HHS Public Access

Author manuscript

Cell Chem Biol. Author manuscript; available in PMC 2018 March 16.

Published in final edited form as:

Cell Chem Biol. 2017 March 16; 24(3): 415–425. doi:10.1016/j.chembiol.2017.02.008.

Directed evolution of a bright near-infrared fluorescent rhodopsin using a synthetic chromophore

Lukas Herwig[a,2], **Austin J. Rice**[a,2], **Claire N. Bedbrook**[a,b], **Ruijie K. Zhang**[a], **Antti Lignell**[a], **Jackson K. B. Cahn**[a], **Hans Renata**[a], **Sheel C. Dodani**[a], **Inha Cho**[a], **Long Cai**[a], **Viviana Gradinaru**[b], and **Frances H. Arnold**[a,b,1]

[a]Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA, 91125

[b]Division of Biology and Biological Engineering, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA, 91125

Summary

By engineering a microbial rhodopsin, Archaerhodopsin-3 (Arch), to bind a synthetic chromophore, merocyanine retinal, in place of the natural chromophore all-*trans*-retinal (ATR), we generated a protein with exceptionally bright and unprecedentedly red-shifted near-infrared (NIR) fluorescence. We show that chromophore substitution generates a fluorescent Arch-complex with a 200 nm bathochromic excitation shift relative to ATR-bound wild-type Arch and an emission maximum at 772 nm. Directed evolution of this complex produced variants with pH-sensitive NIR fluorescence and molecular brightness 8.5-fold greater than the brightest ATR-bound Arch variant. The resulting proteins are well suited to bacterial imaging; expression and stability have not been optimized for mammalian cell imaging. By targeting both the protein and its chromophore we overcome inherent challenges associated with engineering bright NIR fluorescence into Archaerhodopsin. This work demonstrates an efficient strategy for engineering non-natural, tailored properties into microbial opsins, properties relevant for imaging and interrogating biological systems.

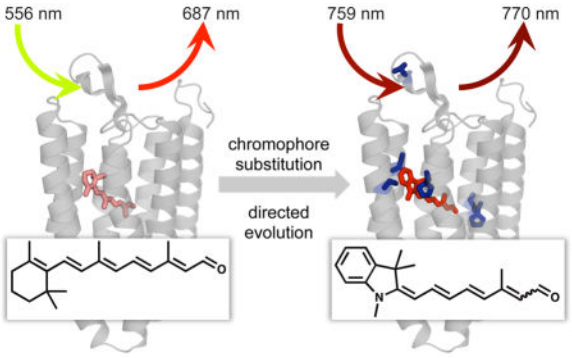
eTOC Blurb

Using a combined approach of chromophore substitution and directed evolution, Herwig *et al.* engineered fluorescent Archaerhodopsin variants with unprecedented near-infrared (NIR) excitation and emission. Evolved variants display pH sensitivity, enhanced fluorescent molecular brightness and improved synthetic chromophore affinity.

[1]To whom correspondence should be addressed. frances@cheme.caltech.edu, Phone: 626-395-4162.
[2]L.H. and A.J.R. contributed equally to this work.

Author Contributions

L.H., A.J.R., and C.N.B. designed and performed research and analyzed data; F.H.A. supervised research; R.K.Z. and H.R. synthesized and characterized merocyanine retinal; A.L. and I.C. performed research; A.L., L.C., and V.G. provided necessary analytical tools; J.K.B.C and S.C.D analyzed data; L.H., A.J.R., C.N.B., and F.H.A. wrote the paper. All authors gave final approval for publication.



Keywords

Synthetic chromophore substitution; directed evolution; protein engineering; near-infrared fluorescence; Archaerhodopsin; live-cell imaging

Introduction

Fluorescent proteins have revolutionized our ability to visualize the biological world at the molecular level (Davidson and Campbell, 2009; Kremers et al., 2011). The photophysical properties of a fluorescent protein are dictated by two factors, the light-excitable chromophore and the interacting protein environment. While nature is very effective at tuning the protein environment through mutation and selection, the limited library of reported natural chromophores ultimately constrains engineering possibilities (Davidson and Campbell, 2009; Shcherbakova and Verkhusha, 2014). Near-infrared (NIR) fluorescence (650–900 nm) is desirable for non-invasive deep tissue imaging due to reduced scattering and low phototoxicity of the longer wavelength light (Weissleder, 2001). Engineered bacterial phytochromes (BphPs) emit NIR fluorescence with their highly conjugated biliverdin chromophore (Fischer and Lagarias, 2004; Shcherbakova et al., 2015b; Shcherbakova and Verkhusha, 2013; Yu et al., 2015), whereas the vitamin A-derived chromophore all-*trans*-retinal (ATR) enables red- to farred fluorescence in certain microbial rhodopsins (Kralj et al., 2011a; Kralj et al., 2011b). Such naturally occurring chromophores are limited in their ability to produce bright, molecular NIR fluorescence. Even with substantial protein engineering efforts, peak fluorescence excitation of ATR-bound rhodopsin variants remain well outside the NIR window (Engqvist et al., 2014; Hochbaum et al., 2014). Among BphPs, there seems to be a trade-off between red shift and molecular brightness, with the brightest variants (brightness equivalent to mCherry (Kremers et al., 2011)) excited outside the NIR (639 nm) and the furthest red-shifted BphPs (peak excitation at 702 nm) approximately 2.7-fold dimmer than mCherry (Shcherbakova et al., 2015a; Shcherbakova and Verkhusha, 2013). Thus, the engineering of bright NIR fluorescent proteins is still an outstanding challenge.

Previous engineering efforts have shown that directed evolution can be used to enhance and red-shift the fluorescence of ATR-bound microbial rhodopsins, including Archaerhodopsin-3 (Arch) (Engqvist et al., 2014; Hochbaum et al., 2014; McIsaac et al., 2014). Originally from the halophilic archaea *Halorubrum sodomense*, wild-type (WT) Arch is a yellow light-driven

proton pump that has dim, far-red fluorescence (Kralj et al., 2011a; Mukohata et al., 1999). In Arch, as with many rhodopsins, the ATR chromophore is covalently bound to a conserved lysine residue via a Schiff base (Ernst et al., 2014). Protonation of this base modulates the spectral properties and isomerization of ATR (Ernst et al., 2014). As a result, Arch fluorescence is sensitive to pH and transmembrane voltage gradients (Maclaurin et al., 2013) and has been used to monitor action potentials when expressed in cultured neurons (Kralj et al., 2011a). Directed evolution of the ATR-bound Arch protein generated variants with brighter and red-shifted fluorescence, enabling fluorescent imaging with lower-power light (Flytzanis et al., 2014; Hochbaum et al., 2014; McIsaac et al., 2014). However, the brightest engineered ATR-bound Arch variant was not pH-sensitive and was at least 12-fold dimmer than mCherry; furthermore its peak excitation (~615 nm) remained outside the NIR window (McIsaac et al., 2014).

In order to access Arch variants with bright NIR emission, we drew inspiration from previous demonstrations that desirable fluorescent protein properties could be obtained by expanding the limited repertoire of naturally known chromophores (Plamont et al., 2016; Tamura and Hamachi, 2014; Yapici et al., 2015). Spectral properties of the natural ATR chromophore (Figure 1*A*, Compound 1) can be modulated by adding electron-withdrawing groups (Gaertner et al., 1981; Hendrickx et al., 1995) or changing the length of the conjugated π-bond system (Albeck et al., 1989; Nielsen, 2009). In particular, retinal analogs with extended conjugation have been shown to red-shift the absorption maxima of rhodopsins up to hundreds of nanometers, well into the NIR window (Asato et al., 1990; Sineshchekov et al., 2012). For example, merocyanine retinal (Figure 1*A*, Compound 2), with extended conjugation relative to ATR, was shown by Hoischen and coworkers to bathochromically shift bacteriorhodopsin absorbance by 187 nm (Hoischen et al., 1997). Other chromophore substitutions have modified the spectral and kinetic characteristics of rhodopsin-based optogenetic tools (AzimiHashemi et al., 2014) and modulated the proton pumping capabilities of proteorhodopsin and *Gloeobacter violaceus* rhodopsin (Ganapathy et al., 2015). Thus chromophore substitution enables rapid introduction of desirable photophysical properties; however, it is unlikely that a wild-type opsin will bind a non-natural chromophore preferentially or that its new spectral properties (e.g. NIR fluorescence) will manifest at optimal levels. We have therefore paired chromophore substitution with directed evolution in order to build on novel capabilities conferred by the synthetic chromophore. Here we transcend the natural fluorescent properties of Arch by evolving the protein around a synthetic chromophore, thereby creating variants with exceptional molecular brightness and unprecedentedly red-shifted NIR fluorescence.

Results and Discussion

Merocyanine retinal enhances photophysical properties of Arch

Chromophore substitution can dramatically modify the inherent properties of a rhodopsin and establish a new platform for achieving desirable features by protein engineering. Given the extensive red-shift conferred on bacteriorhodopsin by merocyanine retinal (Hoischen et al., 1997), we selected this chromophore for substitution in Arch as a first step in developing a bright NIR fluorescent protein. Merocyanine retinal (aldehyde) was synthesized as

described in supplemental methods (Figure S1). Addition of merocyanine retinal (final concentration of 1 μM) to *E. coli* cultures expressing the Arch protein produced an opsin-chromophore complex. A covalent retinal Schiff base (RSB) linkage in the complex was confirmed by spectral comparison of denatured rhodopsin to free merocyanine retinal Schiff base (Figure S2). The capacity of merocyanine retinal to form a RSB was anticipated due to previous observations of a different merocyanine retinal analog (with shorter polyene chain) binding to the retinoic binding protein CRABPII (Yapici et al., 2015). The absorbance and fluorescence excitation of merocyanine-bound Arch were red-shifted by more than 200 nm when compared to ATR-bound Arch (ex/em at 556/687 nm) (McIsaac et al., 2014) in purified protein (Figure 1*B* and Table 1) and in whole cells (Figure S3*A*). This represents a 262 nm 'opsin shift' for Arch-bound merocyanine, which is the difference between the absorption maxima of the free aldehyde chromophore (498 nm) and the newly formed Arch complex (760 nm).

Directed evolution of merocyanine-bound Arch enhances opsin-specific fluorescence

Although capable of doing so, wild-type Arch did not evolve specifically to bind merocyanine retinal or to allow energy escape in the form of fluorescence. Thus, we sought to enhance Arch NIR-fluorescence intensity and molecular brightness by directing the evolution of the Arch protein around this synthetic chromophore. Throughout this work we distinguish between molecular brightness and fluorescence intensity. Although both properties quantify the fluorescence light emitted by a fluorescent protein, the former is a photophysical property of the protein that is determined *in vitro*, while the latter is measured directly via plate reader or microscopy and is influenced by factors such as expression level and imaging conditions. Our directed evolution screen selected for greater NIR fluorescence intensity in *E. coli*.

Random mutagenesis of WT Arch by error-prone PCR and screening 1,700 variants for NIR fluorescence (ex/em 760/785 nm) identified variant Mero-1 with the P60S mutation (Figure 2). Site-saturation mutagenesis of Mero-1 at position G61 (selected from a marginal hit in the error-prone PCR library) led to Mero-2, with mutations P60S and G61L (Figure 2). Mutations that increase fluorescent brightness of microbial rhodopsins (Engqvist et al., 2014; Hochbaum et al., 2014; McIsaac et al., 2014) and increase occupancy of associated states in the photocycle (Maclaurin et al., 2013; Wagner et al., 2013) are known to be located proximal to ATR or the Schiff base. Thus, to guide further evolution, we generated a homology model of merocyanine-bound wild-type Arch based on the known structure of Archaerhodopsin-2 (86% amino acid identity; Figure 3) (Kouyama et al., 2014). Sites located within 5 Å of the indolylidene ring of merocyanine retinal (W148, S151 and P196) or the Schiff base (M30, V59, A63, T99) in the homology model were selected for site-saturation mutagenesis in Mero-2. We identified P196G (Mero-3) and S151A. The latter mutation displayed stronger pigment formation in culture compared to Mero-2, but showed no significant increase in overall fluorescence intensity. Previous work has shown that mutation S151A contributes to a hypsochromic absorbance shift in ATR-bound Arch (Sudo et al., 2013). However, combining P196G and S151A in the Mero-2 background yielded Mero-4 (P60S-G61L-S151A-P196G), which displays greater NIR fluorescence intensity (Figure 2). Upon screening 1,800 variants of Mero-4 in a final round of error-prone PCR

mutagenesis, we identified Mero-6 that has two additional mutations, T80A and G132S. Mero-6 exhibited a 10-fold increase in molecular fluorescence intensity over merocyanine-bound wild-type Arch (Figure 2).

Of the six mutations present in Mero-6 (Table 1), three are predicted to lie in the retinal binding pocket (G132S, S151A, and P196G), two near the Schiff base (P60S and G61L), and one on an extra-cellular loop (T80A) (Figure 3). Mutations near the retinal and Schiff base have been shown to affect the spectral properties of ATR-bound Arch (Hochbaum et al., 2014; McIsaac et al., 2014). Our identification of similar sites (non-identical mutations with the exception of P60S (Hochbaum et al., 2014) and S151A (Sudo et al., 2013)) suggests that non-natural retinal analogs can likewise be spectrally modulated by tuning direct protein-chromophore interactions.

To assess the potential of the merocyanine-bound Arch variants as NIR fluorescent makers in live cells, we acquired NIR images of bacterial and eukaryotic cells expressing Arch-WT and Mero-6. For bacterial imaging, *E. coli* expressing either wild-type Arch or Mero-6 in the presence of merocyanine retinal were readily detected with NIR excitation at 727 nm (Coherent CUBE laser, 32mW power) and detection within 766–854 nm (Figure 4*A*, bottom row); moreover, the evolved Mero-6 variant yielded more intense fluorescence visible with lower contrast (Figure 4*A*, middle row). Opsin expression could be tracked by the fused CFP tag (Figure 4*A*, top row) and, unlike NIR fluorescence, expression was independent of merocyanine retinal addition (Figures 4*A* and S3*B*). Fluorescent puncta in cells expressing wild-type Arch are more pronounced in the absence of merocyanine retinal and may be due to protein instability when the protein is expressed in the absence of chromophore. In support of plate-reader measurements (Figure S3*B*), the mean CFP fluorescence intensity quantified from images reveals that expression was modestly reduced (Figure 4*B*) over the course of evolution, while the absolute and CFP-normalized NIR fluorescence were increased significantly (Figure 4*C* and 4*D*). Both NIR results clearly show the advantage of the evolved Arch variant for bacterial imaging applications. However, comparing the results of Figures 2 and 4*D*, the evolved fluorescence enhancement appears to be 2-fold greater when quantified via microscope analysis as opposed to plate reader, which could be due to the greater light intensity and more sensitive detection involved in microscopy imaging (Maclaurin et al., 2013). Photostability measurements in live *E. coli* indicate that merocyanine-bound wild-type Arch bleaches 79% faster than CFP, while Mero-6 bleaches only 49% faster than CFP (ratio of NIR and CFP exponential decay rates; Table 1 and Figure S4), suggesting that photostability has also increased over the evolutionary trajectory. The ability to track merocyanine-treated, opsin-expressing bacteria in deep tissue could find application in a wide range of biological studies. For example, a NIR fluorescent opsin probe could be useful for tracking bacteria in whole animals or infected tumors (Berlec et al., 2015; Cronin et al., 2012).

For NIR fluorescence in eukaryotic cells, GFP-tagged wild type and Mero-6 constructs were built with optimal codon usage and trafficking signals (Gradinaru et al., 2010) (Figure S5*A*). Representative images and quantitative analysis of transfected human embryonic kidney (HEK 293T) cells show that Mero-6 expresses in those eukaryotic cells, though at considerably lower levels than wild-type Arch and with an increased number of fluorescent

puncta and aggregates that may be due to diminished protein stability or trafficking capacity (Figure S5*B* and *C*). However, normalization of the raw NIR fluorescence (Figure S5*B*, right column and S5*D*) by total opsin expression level (GFP fluorescence, Figure S5*C*) indicated that the molecular fluorescence intensity of merocyanine–bound Mero-6 is significantly greater than that of wild-type protein (Figure S5*E*). These results demonstrate that the enhancement of Arch molecular NIR fluorescence intensity achieved by evolution in bacteria transfers to eukaryotic cells. Given the exceptionally red-shifted fluorescence of the merocyanine-bound Arch variants and their relatively high molecular brightness, merocyanine-bound Arch variants offer promise for imaging applications in eukaryotic cells. The absence of ion pumping for Mero-6 in the presence of ATR or merocyanine retinal (Figure S6) would be beneficial for imaging, where active pumping would perturb the transmembrane voltage and local pH. For eukaryotic applications, however, merocyanine-bound Arch variants would need to be further optimized for expression and stability.

In eukaryotic cells, WT Arch and Mero-6 are marginally sensitive to a voltage step with farred light (Figure S5*F*). In bacteria, the pH sensitivity of merocyanine-bound Arch fluorescence was retained throughout the course of evolution. When validating each Arch variant, *E. coli* NIR fluorescence measurements were taken at three pH values (5, 7 and 9). All selected variants displayed enhanced fluorescence intensity at acidic pHs and dimmer fluorescence at alkaline pH (Figure 2*A*). This bright, pH-sensitive NIR fluorescence could be useful in microbiology applications, for instance to monitor bacterial activity and pH microenvironments in biofilms with dim light (Guo et al., 2013; Hidalgo et al., 2009; Schlafer et al., 2015) or to assess deep-tissue, host-pathogen interactions in disease models (Duhring et al., 2015; Vande Velde et al., 2014).

Evolved NIR fluorescent Arch variants bind merocyanine retinal with greater affinity and lose affinity for ATR

Mutations in the chromophore-binding pocket of Arch modify protein-chromophore interactions and can enhance the molecular fluorescent brightness (Hochbaum et al., 2014; McIsaac et al., 2014) and the affinity of Arch for synthetic merocyanine retinal. Improving the affinity of Arch for merocyanine retinal is important for applications where only lower concentrations of the synthetic chromophore are available or desirable. For example, limited availability could be anticipated when the compound is applied indirectly (e.g. in animal food) or when the natural ATR chromophore is competing for protein binding. Moreover, due to possible toxicity of merocyanine retinal at high concentrations (Figure S8), the ability to use lower concentrations is desirable. To determine whether the merocyanine retinal binding affinity changed during the course of evolution, we first measured how the whole-cell NIR fluorescence depended on the concentration of chromophore added (Figure 5*A*). These results allowed us to calculate the concentration of chromophore required for selected Arch variants to reach half maximal fluorescence after four hours of expression; this chromophore concentration is referred to as the ‘apparent K_d’. In *E. coli*, the apparent K_d for merocyanine retinal decreased 2.5-fold from wild-type Arch to Mero-4 (from 3.25 to 1.33 μM), with the most significant decrease (1.8-fold) observed between WT and Mero-2 (Figure 5*B*). Combined, the two mutations of Mero-2 (P60S and G61L) likely affect the conformation of Helix 2 (Vonheijne, 1991; Wilman et al., 2014), possibly modifying the

protein-Schiff base interaction. Of the two additional mutations in Mero-4, P196G may affect the conformation of Helix 4, and both P196G and S151A could expand the retinal-binding pocket to better accommodate the indolylidene ring of merocyanine retinal. Accurate structural information will be required for a more detailed analysis of the merocyanine conformation and interacting opsin residues.

As a more direct measurement of the affinity of Arch for ATR and merocyanine, we also monitored the absorbance of purified, *apo* opsin pigmented with increasing concentrations of each chromophore. Due to differences in binding pocket accessibility of the detergent-stabilized *apo* protein versus the unpurified membrane-bound opsins (and the likelihood of co-translational pigmentation in the latter), the absolute binding affinities determined by the two methods are not expected to be the same. However, the relative affinities between Arch variants in *E. coli* (Figure 5*B*) and purified protein (Table 1) show a similar trend of increasing affinity for the new chromophore, particularly in the initial rounds of evolution. For purified protein, the K_d of merocyanine retinal decreased from 99 μM (wild-type Arch) to 11 μM (Mero-2) and increased slightly in the final rounds of evolution to 17 μM (Mero-6). Therefore, the *in vitro* results indicate that mutations made in in the initial rounds of evolution enhanced NIR fluorescence as well as merocyanine affinity. Moreover, the affinity for ATR decreased during the course of evolution, as the K_d increased from 8 μM in wild-type Arch to 330 μM in Mero-6. Possible differences in the detergent partition coefficients of merocyanine retinal and ATR prevent direct comparison of calculated binding affinities of the two lipophilic chromophores; however, the affinities of different Arch variants for a given chromophore can be compared. The opposite trends in merocyanine retinal and ATR binding affinity indicate that, over the course of evolution, the binding pocket was restructured to accommodate the synthetic chromophore at the expense of ATR.

Interestingly, the last round of evolution shows no further improvement in merocyanine binding affinity despite a 50% increase in fluorescence between Mero-4 and Mero-6 in *E. coli* (Figure 5*B*). Thus enhanced affinity played an important role in the early stages of Arch evolution, but the increased *E. coli* fluorescence observed in the later rounds reflects a change in the photophysical properties of merocyanine-bound opsin.

Spectral properties of evolved merocyanine retinal-bound Arch variants

Merocyanine-bound Arch variants were purified and their spectral properties were characterized. For all tested variants (Figure S7), the peak excitation and emission values were well within the NIR window (650–900 nm (Weissleder, 2001)). Consistent with other red-shifted fluorescent proteins (Shcherbakova et al., 2012), the merocyanine-bound opsins have a small Stokes shift (11–17 nm). However, a long emission tail extending past 800 nm allows for fluorescence detection that is deeper into the NIR region and is well separated from peak excitation.

The photophysical characteristics of the merocyanine-bound variants were found to have improved over the course of directed evolution. The quantum yield (QY), extinction coefficient (ε), and molecular brightness ((QY × ε) / 1000) all increased slightly from wild-type Arch to Mero-2, then decreased slightly between Mero-2 and Mero-4 (Table 1). Between Mero-4 and Mero-6, the extinction coefficient remained the same, but the QY and

therefore brightness increased significantly (~24% increase over Mero-4). Thus the two new mutations in Mero-6 (G132S and T80A) do not affect the affinity for merocyanine (Figure 5*B*), but do enhance the QY and brightness of the merocyanine-Arch complex. The molecular basis of this QY enhancement is difficult to pinpoint without detailed structural information. However, a T80S mutation was shown to contribute to the fluorescence of ATR-bound Arch in a previous study (Hochbaum et al., 2014), and G132S introduces a polar group near the indolylidene ring, potentially stabilizing electron delocalization. Mutation G132S alters a protein residue not previously known to increase Arch fluorescence. However, a G132V mutation was involved in converting Arch from a light-driven proton pump into a light-gated proton channel (Inoue et al., 2015).

To assess the spectral properties of these Arch variants evolved around a synthetic chromophore, we drew comparisons to other fluorescent proteins. Compared to the exceptionally dim fluorescence of wild-type Arch bound to native ATR, merocyanine-bound Mero-6 is 16- to 200-fold brighter. The range reflects the fact that quantum yield depends on light intensity, which makes ATR-bound wild-type Arch brighter at high intensity light (Maclaurin et al., 2013) and dimmer at low intensity light (Kralj et al., 2011a). An engineered variant of Arch with seven mutations, termed ‘Arch-7,’ had the highest molecular brightness of any ATR-bound Arch and was also the furthest red-shifted, with an excitation peak at 615 nm (McIsaac et al., 2014). Compared to Arch-7, merocyanine-bound Mero-6 is 8.5 times brighter and 143 nm further red-shifted, reaching an unprecedented peak excitation wavelength of 759 nm. Mero-6 has about 68% the molecular brightness of the soluble red-fluorescent protein mCherry (Kremers et al., 2011), but is excited by light 172 nm further red-shifted. Bacterial phytochrome photoreceptors (BphPs) have been developed as NIR-fluorescent cellular markers (Shcherbakova et al., 2015b). The brightest BphP is 50% brighter than Mero-6; however, Mero-6 is excited by light 120 nm further red (Shcherbakova et al., 2015a). Compared to the most red-shifted BphP, Mero-6 is excited 57 nm further into the NIR and is 84% brighter (Shcherbakova and Verkhusha, 2013). These results demonstrate that chromophore substitution in microbial rhodopsins using a tailored chromophore and subsequent directed evolution is a powerful and effective way to engineer desired fluorescent protein properties.

Concluding Remarks

Customized fluorescent proteins can be generated by engineering naturally occurring scaffolds to bind synthetic chromophores (Paige et al., 2011; Plamont et al., 2016; Tamura and Hamachi, 2014; Yapici et al., 2015). Chromophore-dependent microbial opsins provide an excellent platform for this approach since modified retinals are well accepted and their incorporation can dramatically alter and enhance opsin properties (Albeck et al., 1989; Asato et al., 1990; AzimiHashemi et al., 2014; Gaertner et al., 1981; Ganapathy et al., 2015; Hoischen et al., 1997; Nielsen, 2009; Sineshchekov et al., 2012). We have used a synthetic chromophore, merocyanine retinal, to generate a near-infrared fluorescent protein that is 200 nm red-shifted and at least 16-fold brighter than WT Arch bound to its natural chromophore, ATR. Directed evolution of Arch around the highly conjugated merocyanine retinal allowed us to enhance chromophore affinity, diminish affinity for the native retinal, and further augment pH-sensitive fluorescent brightness specifically in the NIR window. The Arch

variants were engineered in *E. coli* and are well suited to bacterial imaging applications. Although we found that merocyanine-dependent NIR fluorescence could be detected in eukaryotic (HEK) cells, expression and stability of Arch should be optimized for eukaryotic imaging applications. With appropriate screens, one could use directed evolution to enhance other properties of interest such as enhanced expression in specific cell types, high voltage sensitivity, or pumping activity using tailored synthetic retinals that are sensitive to specific wavelengths. We anticipate that this general approach will continue to provide fine-tuned properties and functions useful for optogenetic sensors or actuators (AzimiHashemi et al., 2014; Sineshchekov et al., 2012), energy harvesting (Ganapathy et al., 2015; Walter et al., 2010), or cell labeling and imaging applications (Albeck et al., 1989; Hoischen et al., 1997).

Significance

At the core of every fluorescent protein, a chromophore absorbs light and emits it as fluorescence. The molecular brightness and color of this fluorescence are controlled by the chromophore and surrounding protein environment. Near-infrared (NIR) fluorescent proteins are desirable for deep tissue imaging yet their development represents a significant engineering challenge. Here, we engineer both protein and chromophore in Archaerhodopsin-3 (Arch) to access fluorescent properties in the near-infrared region. By evolving Arch around a synthetic chromophore, we obtained a near-infrared fluorescent rhodopsin with exceptional molecular brightness. This engineering strategy provides an efficient route to develop rhodopsin complexes with properties relevant for optogenetics, energy harvesting, or *in vivo* imaging applications.

Experimental Procedures

Synthesis and characterization of merocyanine retinal

In brief, merocyanine retinal was synthesized (Figure S1) as described (Hoischen et al., 1997) and characterized by NMR and high-resolution mass spectrometry (Data S1). Full details of the synthesis are provided in the supplemental information.

Cloning, plasmids and bacterial strains

An *Escherichia coli* codon-optimized version of 6xHis-tagged wild-type Arch (g-block from Integrated DNA Technologies (IDT)) was used for mutant library construction and fluorescence screening. Mutations identified in the 6xHis-tagged wild-type Arch construct were transferred to the previously described Arch construct (pETME14-CFP) (McIsaac et al., 2014) for Arch fluorescence normalization (wild-type sequences are given in the SI). For cloning and directed evolution, we used electroporation and *E. cloni*® EXPRESS BL-21(DE3) cells (Lucigen). Chemically competent NiCo pLEMO cells (NEB) were used for large-scale expression.

Directed evolution

Library construction—Random mutagenesis libraries were generated via error-prone PCR with a range of $MnCl_2$ concentrations (100 – 600 μM $MnCl_2$). For each PCR reaction, we used a final concentration of [0.08 Units/μL Taq polymerase; 400 μM dNTPs; 1× Taq

standard buffer; 0.2 μM of each primer] in 100 μL total reaction volume. The error-prone PCR thermocycler conditions were as follows: 2 min at 95 °C (1×); 30 s at 95 °C, 30 s at 55 °C, 1 min at 72 °C (30×); 10 min at 72 °C (1×),10 °C ∞. This resulted in average library error rates of 2–3 nucleotides per gene. The mutated Arch PCR products were cloned into the pET21a expression plasmid (EMD Millipore) by isothermal assembly (Gibson et al., 2009). Ultimately, 1,700 colonies from the 500 μM $MnCl_2$ library on wild-type Arch and 1,800 colonies from the 600 μM $MnCl_2$ library on Mero-4 were screened for increased opsin fluorescence as described below. Site-saturation mutagenesis libraries were generated by amplifying the parental plasmid DNA (Mero-2) with mutagenic NNK primers as described previously (Engqvist et al., 2014).

Expression and screening of mutant Arch libraries—Screening of the error-prone PCR and site-saturation libraries was done using a 6xHis-tagged version of Arch at pH 7. Putative improved variants were re-screened in sextuples at pH 5, pH 7, and pH 9. Mutations were identified by Sanger sequencing (Laragen). Single colonies from libraries were selected with sterile toothpicks and inoculated into 300 μL of Luria broth (LB; 100 μg/mL carbenicillin) in sterile deep-well 96-well plates. Plates were sealed with EasyApp microporous films (part no. 2977–6202; USA Scientific, Inc.). Following overnight growth at 37 °C, 220 rpm and 80% humidity, pre-cultures were diluted 1:20 into 1 mL of fresh LB (100 μg/mL carbenicillin) and grown for 2 h at 30 °C. Then merocyanine retinal (1 μM final concentration) and isopropyl β-D-1 thiogalactopyranoside (IPTG; 500 μM final concentration) were added to each well, and proteins were expressed for 4–6 h in the dark. Next, cells were centrifuged at 4,000 RPM in a swinging bucket rotor (Beckman Coulter, Allegra TM 25R Centrifuge). The resulting pellets were resuspended in 700 μL of 200 mM NaCl, and 180 μL of the resuspension was added to 20 μL of 500 mM potassium phosphate buffer (at pH 7 for the initial screen and pH 5, 7 and 9 in the re-screen) in a measurement plate (Evergreen Scientific, untreated 96-well microplates, catalog number: 290-8115-01F). Raw Arch fluorescence was measured using a Tecan Infinite® M200 plate-reader at an emission wavelength of 785 nm following excitation at 760 nm. The mutations identified were transferred into the pETME14-CFP construct (McIsaac et al., 2014). The C-terminal cyan fluorescent protein (CFP; ex/em 425/475 nm) tag was used as a proxy for expression level and to calculate normalized opsin fluorescence. The reported normalized fluorescence is defined as 1000 × [opsin fluorescence/CFP fluorescence]. Screening of site-saturation libraries was performed as described above. We analyzed 88 clones for each site, for 94% coverage of the possible diversity (Engqvist et al., 2014). Data from all libraries were processed and analyzed using Microsoft Excel and GraphPad Prism (version 6.04 for Windows; GraphPad Software) software. Cells expressing a non-fluorescent protein ScADH6 (*Saccharomyces cerevisiae* cinnamyl alcohol dehydrogenase) were used as the negative control in all rounds of directed evolution.

Arch homology model

A homology model, built using the Swiss modeler web server (Arnold et al., 2006; Bordoli et al., 2008), was used to identify residues of interest for site-saturation mutagenesis (Figure 3). A three-dimensional model of merocyanine retinal was generated using the NCI CACTUS server (http://cactus.nci.nih.gov/translate/) and manually inserted into the

homology model via Coot (Emsley and Cowtan, 2004) and the UCSF Chimera package (Pettersen et al., 2004).

Purification of Arch variants

Arch variants of interest were expressed and purified as fully described in the SI. In brief, 1L cultures of Arch variants were grown at 37°C, induc ed at mid-log growth by adding merocyanine retinal and IPTG (final concentrations of 1 μM and 0.5 mM, respectively), grown post-induction for 3 hours at 30°C in the dark, the n harvested and stored at −80 °C. *Holo*-Arch variants were purified as reported (McIsaac et al., 2014) via Ni-affinity chromatography. Arch-containing elution fractions were pooled and immediately desalted into 'DDM desalt buffer' [20 mM Tris-HCl; pH 6.5; 200 mM NaCl; 0.15% DDM] via PD-10 desalting columns.

For *in vitro* chromophore binding studies, *apo*-Arch variants were obtained from cultures induced in the absence of chromophore. All steps were carried out at 4 °C. Thawed cells were lysed via microfluization in the absence of detergent. A crude membrane fraction was collected via ultra-centrifugation, stored at −80 °C, and sol ubilized with 1.5% lauryl maltose neopentyl glycol (LMNG) detergent. Solubilized *apo*-Arch was then purified via Ni-affinity chromatography. Arch containing fractions were identified via SDS-PAGE and desalted into [20 mM Tris-HCl; pH 7.5; 200 mM NaCl; 0.015% LMNG] ('LMNG desalt buffer') via PD-10 desalting columns. For the binding assay, purified proteins were concentrated no more than 2× via spin filtration (Millipore).

K_d measurements in *E. coli* and of purified proteins

For the *E. coli* K_d measurements, protein expression and fluorescence measurements were performed as described in the directed evolution methods. For each retinal concentration and variant, the fluorescence emission was recorded for the opsin (ex/em 760/785 nm) and CFP (ex/em at 425/475 nm) after 4 hours of post-induction incubation with merocyanine retinal. Each variant was tested under the following merocyanine retinal concentrations: 32, 16, 4, 1, 0.25, 0.125, 0.063 and 0 μM. The use of an isomeric mixture of merocyanine retinal (3:1 *trans*: 16-*cis*) may lead to underestimated chromophore binding affinities; however, trends in total merocyanine retinal binding affinity can still be assessed. Normalized emission intensity was calculated in quadruplicate and the average with standard deviation was plotted versus the retinal concentration. The data were fitted to the following equation $F = (F_{max}[retinal]/(K_d + [retinal])$, where F is the observed fluorescence and F_{max} is the calculated maximal fluorescence of the retinal-opsin complex. K_d values of the selected Arch variants were determined from the generated plot using GraphPad Prism (version 6.04 for Windows; GraphPad Software) software. Co-translational chromophore binding in this assay would lead to variable on- and off-rates during the course of protein maturation, which would complicate the assumption of equilibrium. So the K_d values measured in this assay are referred to as 'apparent K_d.'

In vitro, the binding of ATR or merocyanine retinal to *apo*-Arch was monitored by the 'opsin shift' in chromophore absorbance (Baloghnair et al., 1981; Booth et al., 1996). To measure the binding affinity, purified protein, LMNG desalt buffer, and finally retinal dissolved in

ethanol were added to a 96-well half-area plate for a final volume of 100 μL (final concentration of ethanol < 1%). With a constant concentration of protein, a range of retinal concentrations was used (0, 0.1, 0.25, 0.5, 0.75, 1, 2, 5, and 10 molar equivalents of retinal to protein) so that a binding curve could be determined. The final protein concentration was set to 0.508 mg/mL for wild-type Arch, Mero-2, and Mero-4 (0.432 mg/mL for Mero-6), which would allow accurate absorbance measurements of the *holo* protein. At these protein concentrations, more than a 10:1 molar equivalent of merocyanine retinal to protein resulted in protein aggregation. For each Arch variant and concentration, the binding affinity of ATR and merocyanine retinal were measured in triplicate. K_d values calculated from data measured after 24 and 36 hours of incubation remain consistent (average percent difference of 9%) indicating that binding has reached equilibrium by 36 hours; thus the 36 hour data were used to calculate binding affinities in Table 1. Using a custom fitting function in Matlab (MATLAB 8.5, The MathWorks Inc., 2015), curves were fitted to the data without the common assumption that added retinal approximates free retinal (i.e. accounting for ligand depletion) (Swillens, 1995). For this fitting scheme, *holo* protein absorbance was converted to concentration via Beer's law; the path-length for this volume of sample was calculated at 0.6636 cm, the extinction coefficients for merocyanine-bound Arch variants are determined below, the extinction coefficient for ATR-bound Arch was reported previously (McIsaac et al., 2014), and the extinction coefficient for ATR bound Mero-6 was approximated by the wild-type value. We found that the K_d parameter of the fit was only marginally influenced by changes in extinction coefficient.

Quantum yield and extinction coefficient determination

Spectral properties of purified, merocyanine-bound variants were characterized as previously described (McIsaac et al., 2014; Wall et al., 2015) and detailed in the SI. In brief, emission spectra (Figure S7A) were collected via plate reader with excitation at 700 nm and emission detected between 730 and 850 nm; excitation spectra (Figure S7B) were collected with detection at 810 nm and excitation scanned from 690 to 790 nm. As shown in Figure S9A–B, the quantum yield for each merocyanine-bound Arch variant was calculated by comparison to the to the Alexa Fluor® 750 NHS Ester (succinimidyl ester) dye with known quantum yield of 0.12 (catalog number: A20011, Life Technologies Corporation) (Wurth et al., 2012). The extinction coefficient of free merocyanine retinal (oxime) was determined via dilution and absorbance measurements. The extinction coefficient of each merocyanine-bound Arch variant was determined by comparing changes in free and bound chromophore while chemically bleaching the *in vitro* sample (Figure 9C–F).

Cell maintenance for live cell imaging

Starting from an overnight pre-culture (LB + 100 μg/mL carbenicillin), 5 mL *E. coli* cultures were grown to early log phase (0.4–0.6 OD_{600}) at 37 °C, then induced with 1 μM merocyanine retinal and 0.5 mM IPTG (final concentrations) at 30 °C for 3 hours. Cells were harvested via centrifugation and stored on ice. Within three hours, the cells were resuspended in 1× PBS and aliquoted on freshly prepared agarose pads for imaging.

As fully described in the supplemental methods, HEK 293T cells were grown in D10 medium (Dulbecco's modified Eagle medium (DMEM) supplemented with 10% (vol/vol)

FBS, 1% sodium bicarbonate and 1% sodium pyruvate) and transfected with GFP-tagged Arch variants that had been codon optimized for mammalian expression. Twenty-four hours after transfection, the D10 medium was replaced with D10 medium supplemented with 1 μM merocyanine retinal. Imaging was done 48 h post transfection. After this prolonged incubation, HEK cells were adherent with normal morphology and healthy appearance (Figure S8).

E. coli and HEK cell imaging and data processing

As fully described in the supplemental methods, cells were imaged with two orthogonal channels (405 nm and 727 nm for *E. coli* or 473 nm and 727 nm for HEK) with corresponding filter cube sets (detection at >418 nm, >498 nm, and 766–854 nm for laser illumination at 405 nm, 473 nm, and 727 nm respectively). Two oil objectives, Olympus NA 1.40 UPlanSApo 100× with additional 1.6× magnification for *E. coli* and Olympus NA 1.35 UPlanSApo 60× for HEK cells were used for imaging. The camera was back-illuminated CCD Andor iKon-M 934 BEX2-DD, offering high quantum efficiency in the near infrared region with a pixel size of 13×13 μm. Fluorescence analysis of *E. coli* clusters was done by masking the background in the CFP image (via an otsu threshold of pixel intensity counts) and determining the mean pixel intensity within the mask (signal) and outside the mask (background). The background-corrected mean pixel intensities are reported in Figure 4*B–C*. The NIR/CFP ratio was determined for each cluster of cells and the ratios were averaged for the values given in Figure 4*D*. Fluorescence analysis of single HEK cells was done by manually selecting regions around each cell and separately a background region in open source ImageJ (version v1.48). Mean fluorescence intensity measurements were recorded for each region of interest (ROI). Background mean intensity was then used to background subtract from the cell mean intensity.

Fluorescence photo-bleaching in live *E. coli*

As fully described in the SI, fluorescence decay rates were measured for both CFP (111 mW, ex/em at 405/464–500 nm, 500 ms exposure) and merocyanine-bound Arch (32 mW ex/em at 727/770–840 nm, 250 ms exposure) allowing Arch decay rates to be presented relative to CFP. The fluorescence decay curves (between 3 and 60 s) were fit with a single exponential using the scipy.stats module in python. The [NIR fluorescence / CFP fluorescence] ratio of exponential decay rates was calculated for each spot of clustered *E. coli* cells. The mean ratio for a given Arch variant (n = 4–5 spots) is given in Table 1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

L.H. was supported by a fellowship from the SNSF (Swiss National Science Foundation). The Ruth L. Kirschstein National Research Service Award supports A.J.R (F32GM116319), C.N.B (F31MH102913), and S.C.D (5F32GM106618). R.K.Z. was supported by a National Science Foundation Graduate Research Fellowship (NSF GRFP; DGE-1144469), is a trainee in the Caltech Biotechnology Leadership Program, and has received financial support from the Donna and Benjamin M. Rosen Bioengineering Center. J.K.B.C. acknowledges the support of the

Resnick Sustainability Institute (Caltech). Research is supported by the National Center For Research Resources, ARRA SIG Program S10RR027203 (F.H.A); National Institute for Mental Health R21MH103824 (V.G. and F.H.A); and the Institute for Collaborative Biotechnologies through grant number W911F-09-0001 from the U.S. Army Research Office (F.H.A). The authors would like to acknowledge the Beckman Institute for the Resource Center on CLARITY, Optogenetics, and Vector Engineering for technology development and broad dissemination (http://www.beckmaninstitute.caltech.edu/clover.shtml). The content is solely the responsibility of the authors and does not necessarily reflect the position or policy of the National Center for Research Resources, the National Institutes of Health, or the Government, and no official endorsement should be inferred.

The authors would like to thank Sabine Brinkmann-Chen and Andrew Buller for critical review of the manuscript and David VanderVelde for assistance with NMR analysis.

References

Albeck A, Friedman N, Sheves M, Ottolenghi M. Factors affecting the absorption maxima of acidic forms of bacteriorhodopsin – A study with artificial pigments. Biophysical journal. 1989; 56:1259–1265. [PubMed: 2611336]

Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006; 22:195–201. [PubMed: 16301204]

Asato AE, Li XY, Mead D, Patterson GML, Liu RSH. Azulenic retinoids and the corresponding bacteriorhodopsin analogs. Unusually red-shifted pigments. Journal of the American Chemical Society. 1990; 112:7398–7399.

AzimiHashemi N, Erbguth K, Vogt A, Riemensperger T, Rauch E, Woodmansee D, Nagpal J, Brauner M, Sheves M, Fiala A, et al. Synthetic retinal analogues modify the spectral and kinetic characteristics of microbial rhodopsin optogenetic tools. Nature Communications. 2014; 5:5810.

Baloghnair V, Carriker JD, Honig B, Kamat V, Motto MG, Nakanishi K, Sen R, Sheves M, Tanis MA, Tsujimoto K. The opsin shift in bacteriorhodopsin – Studies with artificial bacteriorhodopsins. Photochemistry and Photobiology. 1981; 33:483–488.

Berlec A, Zavrsnik J, Butinar M, Turk B, Strukelj B. In vivo imaging of Lactococcus lactis, Lactobacillus plantarum and Escherichia coli expressing infrared fluorescent protein in mice. Microbial Cell Factories. 2015; 14

Booth PJ, Farooq A, Flitsch SL. Retinal binding during folding and assembly of the membrane protein bacteriorhodopsin. Biochemistry. 1996; 35:5902–5909. [PubMed: 8639552]

Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nature Protocols. 2008; 4:1–13.

Cronin M, Akin AR, Collins SA, Meganck J, Kim J-B, Baban CK, Joyce SA, van Dam GM, Zhang N, van Sinderen D, et al. High Resolution In Vivo Bioluminescent Imaging for the Study of Bacterial Tumour Targeting. Plos One. 2012; 7

Davidson MW, Campbell RE. Engineered fluorescent proteins: innovations and applications. Nature Methods. 2009; 6:713–717. [PubMed: 19953681]

Duhring S, Germerodt S, Skerka C, Zipfel PF, Dandekar T, Schuster S. Host-pathogen interactions between the human innate immune system and Candida albicans - understanding and modeling defense and evasion strategies. Front Microbiol. 2015; 6

Emsley P, Cowtan K. *Coot* : model-building tools for molecular graphics. Acta Crystallographica Section D Biological Crystallography. 2004; 60:2126–2132. [PubMed: 15572765]

Engqvist MKM, McIsaac RS, Dollinger P, Flytzanis NC, Abrams M, Schor S, Arnold FH. Directed Evolution of Gloeobacter violaceus Rhodopsin Spectral Properties. Journal of molecular biology. 2014

Ernst OP, Lodowski DT, Elstner M, Hegemann P, Brown LS, Kandori H. Microbial and Animal Rhodopsins: Structures, Functions, and Molecular Mechanisms. Chemical Reviews. 2014; 114:126–163. [PubMed: 24364740]

Fischer AJ, Lagarias JC. Harnessing phytochrome's glowing potential. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101:17334–17339. [PubMed: 15548612]

Flytzanis NC, Bedbrook CN, Chiu H, Engqvist MKM, Xiao C, Chan KY, Sternberg PW, Arnold FH, Gradinaru V. Archaerhodopsin variants with enhanced voltage-sensitive fluorescence in mammalian and Caenorhabditis elegans neurons. Nature Communications. 2014; 5:4894.

Gaertner W, Oesterhelt D, Towner P, Hopf H, Ernst L. 13-(Trifluoromethyl) retinal forms an active and far-red-shifted chromophore in bacteriorhodopsin. Journal of the American Chemical Society. 1981; 103:7642–7643.

Ganapathy S, Bécheau O, Venselaar H, Frölich S, van der Steen Jeroen B, Chen Q, Radwan S, Lugtenburg J, Hellingwerf Klaas J, de Groot Huub JM, et al. Modulation of spectral properties and pump activity of proteorhodopsins by retinal analogues. Biochemical Journal. 2015; 467:333–343. [PubMed: 25655771]

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods. 2009; 6:343–345. [PubMed: 19363495]

Gradinaru V, Zhang F, Ramakrishnan C, Mattis J, Prakash R, Diester I, Goshen I, Thompson KR, Deisseroth K. Molecular and Cellular Approaches for Diversifying and Extending Optogenetics. Cell. 2010; 141:154–165. [PubMed: 20303157]

Guo L, Hu W, He X, Lux R, McLean J, Shi W. Investigating Acid Production by Streptococcus mutans with a Surface-Displayed pH-Sensitive Green Fluorescent Protein. Plos One. 2013; 8

Hendrickx E, Clays K, Persoons A, Dehu C, Bredas JL. The Bacteriorhodopsin Chromophore Retinal and Derivatives - an Experimental and Theoretical Investigation of the 2nd-Order Optical-Properties. Journal of the American Chemical Society. 1995; 117:3547–3555.

Hidalgo G, Burns A, Herz E, Hay AG, Houston PL, Wiesner U, Lion LW. Functional Tomographic Fluorescence Imaging of pH Microenvironments in Microbial Biofilms by Use of Silica Nanoparticle Sensors. Applied and Environmental Microbiology. 2009; 75:7426–7435. [PubMed: 19801466]

Hochbaum DR, Zhao Y, Farhi SL, Klapoetke N, Werley CA, Kapoor V, Zou P, Kralj JM, Maclaurin D, Smedemark-Margulies N, et al. All-optical electrophysiology in mammalian neurons using engineered microbial rhodopsins. Nature Methods. 2014; 11:825–833. [PubMed: 24952910]

Hoischen D, Gärtner W, Buss V, Steinmüller S, Martin HD. Merocyanines as extremely bathochromically absorbing chromophores in the halobacterial membrane protein bacteriorhodopsin. Angewandte Chemie International Edition in English. 1997; 36:1630–1633.

Inagaki HK, Jung Y, Hoopfer ED, Wong AM, Mishra N, Lin JY, Tsien RY, Anderson DJ. Optogenetic control of Drosophila using a red-shifted channelrhodopsin reveals experience-dependent influences on courtship. Nature Methods. 2014; 11:325–U311. [PubMed: 24363022]

Inoue K, Tsukamoto T, Shimono K, Suzuki Y, Miyauchi S, Hayashi S, Kandori H, Sudo Y. Converting a Light-Driven Proton Pump into a Light-Gated Proton Channel. Journal of the American Chemical Society. 2015; 137:3291–3299. [PubMed: 25712566]

Kouyama T, Fujii R, Kanada S, Nakanishi T, Chan SK, Murakami M. Structure of archaerhodopsin-2 at 1.8 angstrom resolution. Acta Crystallographica Section D-Biological Crystallography. 2014; 70:2692–2701.

Kralj JM, Douglass AD, Hochbaum DR, Maclaurin D, Cohen AE. Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. Nature Methods. 2011a; 9:90–95. [PubMed: 22120467]

Kralj JM, Hochbaum DR, Douglass AD, Cohen AE. Electrical Spiking in Escherichia coli Probed with a Fluorescent Voltage-Indicating Protein. Science. 2011b; 333:345–348. [PubMed: 21764748]

Kremers GJ, Gilbert SG, Cranfill PJ, Davidson MW, Piston DW. Fluorescent proteins at a glance. Journal of Cell Science. 2011; 124:2676–2676.

Maclaurin D, Venkatachalam V, Lee H, Cohen AE. Mechanism of voltage-sensitive fluorescence in a microbial rhodopsin. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110:5939–5944. [PubMed: 23530193]

McIsaac RS, Engqvist MKM, Wannier T, Rosenthal AZ, Herwig L, Flytzanis NC, Imasheva ES, Lanyi JK, Balashov SP, Gradinaru V, et al. Directed evolution of a far-red fluorescent rhodopsin. Proceedings of the National Academy of Sciences. 2014; 111:13034–13039.

Mukohata Y, Ihara K, Tamura T, Sugiyama Y. Halobacterial rhodopsins. Journal of Biochemistry. 1999; 125:649–657. [PubMed: 10101275]

Nielsen MB. Model systems for understanding absorption tuning by opsin proteins. Chemical Society Reviews. 2009; 38:913. [PubMed: 19421571]

Paige JS, Wu KY, Jaffrey SR. RNA mimics of green fluorescent protein. Science. 2011; 333:642–646. [PubMed: 21798953]

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF chimera - A visualization system for exploratory research and analysis. Journal of Computational Chemistry. 2004; 25:1605–1612. [PubMed: 15264254]

Plamont MA, Billon-Denis E, Maurin S, Gauron C, Pimenta FM, Specht CG, Shi J, Querard J, Pan B, Rossignol J, et al. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo. Proceedings of the National Academy of Sciences of the United States of America. 2016; 113:497–502. [PubMed: 26711992]

Schlafer S, Garcia JE, Greve M, Raarup MK, Nyvad B, Dige I. Ratiometric Imaging of Extracellular pH in Bacterial Biofilms with C-SNARF-4. Applied and Environmental Microbiology. 2015; 81:1267–1273. [PubMed: 25501477]

Shcherbakova DM, Baloban M, Pletnev S, Malashkevich VN, Xiao H, Dauter Z, Verkhusha VV. Molecular Basis of Spectral Diversity in Near-Infrared Phytochrome-Based Fluorescent Proteins. Chemistry & Biology. 2015a; 22:1540–1551. [PubMed: 26590639]

Shcherbakova DM, Baloban M, Verkhusha VV. Near-infrared fluorescent proteins engineered from bacterial phytochromes. Current Opinion in Chemical Biology. 2015b; 27:52–63. [PubMed: 26115447]

Shcherbakova DM, Subach OM, Verkhusha VV. Red Fluorescent Proteins: Advanced Imaging Applications and Future Design. Angewandte Chemie-International Edition. 2012; 51:10724–10738. [PubMed: 22851529]

Shcherbakova DM, Verkhusha VV. Near-infrared fluorescent proteins for multicolor in vivo imaging. Nature Methods. 2013; 10:751–754. [PubMed: 23770755]

Shcherbakova DM, Verkhusha VV. Chromophore chemistry of fluorescent proteins controlled by light. Current Opinion in Chemical Biology. 2014; 20:60–68. [PubMed: 24819887]

Sineshchekov OA, Govorunova EG, Wang J, Spudich JL. Enhancement of the Long-Wavelength Sensitivity of Optogenetic Microbial Rhodopsins by 3,4-Dehydroretinal. Biochemistry. 2012; 51:4499–4506. [PubMed: 22577956]

Sudo Y, Okazaki A, Ono H, Yagasaki J, Sugo S, Kamiya M, Reissig L, Inoue K, Ihara K, Kandori H, et al. A Blue-shifted Light-driven Proton Pump for Neural Silencing. Journal of Biological Chemistry. 2013; 288:20624–20632. [PubMed: 23720753]

Swillens S. Interpretation of binding curves obtained with high receptor concentrations - practical aid for computer-analysis. Molecular Pharmacology. 1995; 47:1197–1203. [PubMed: 7603460]

Tamura T, Hamachi I. Recent Progress in Design of Protein-Based Fluorescent Biosensors and Their Cellular Applications. Acs Chemical Biology. 2014; 9:2708–2717. [PubMed: 25317665]

Vande Velde G, Kucharikova S, Schrevens S, Himmelreich U, Van Dijck P. Towards non-invasive monitoring of pathogen-host interactions during Candida albicans biofilm formation using in vivo bioluminescence. Cell Microbiol. 2014; 16:115–130. [PubMed: 23962311]

Vonheijne G. Proline kinks in transmembrane alpha-helicies. Journal of molecular biology. 1991; 218:499–503. [PubMed: 2016741]

Wagner NL, Greco JA, Ranaghan MJ, Birge RR. Directed evolution of bacteriorhodopsin for applications in bioelectronics. J R Soc Interface. 2013; 10:20130197. [PubMed: 23676894]

Wall KP, Dillon R, Knowles MK. Fluorescence Quantum Yield Measurements of Fluorescent Proteins: A Laboratory Experiment for a Biochemistry or Molecular Biophysics Laboratory Course. Biochemistry and Molecular Biology Education. 2015; 43:52–59. [PubMed: 25395254]

Walter JM, Greenfield D, Liphardt J. Potential of light-harvesting proton pumps for bioenergy applications. Current Opinion in Biotechnology. 2010; 21:265–270. [PubMed: 20371172]

Wang Y, Zhou M, Wang X, Qin G, Weintraub NL, Tang Y. Assessing in vitro stem-cell function and tracking engraftment of stem cells in ischaemic hearts by using novel iRFP gene labelling. Journal of Cellular and Molecular Medicine. 2014; 18:1889–1894. [PubMed: 24912616]

Weissleder R. A clearer vision for in vivo imaging. Nature biotechnology. 2001; 19:316–316.

Wilman HR, Shi J, Deane CM. Helix kinks are equally prevalent in soluble and membrane proteins. Proteins-Structure Function and Bioinformatics. 2014; 82:1960–1970.

Wurth C, Pauli J, Lochmann C, Spieles M, Resch-Genger U. Integrating Sphere Setup for the Traceable Measurement of Absolute Photoluminescence Quantum Yields in the Near Infrared. Anal Chem. 2012; 84:1345–1352. [PubMed: 22242570]

Yapici I, Lee KSS, Berbasova T, Nosrati M, Jia X, Vasileiou C, Wang W, Santos EM, Geiger JH, Borhan B. "Turn-On" Protein Fluorescence: In Situ Formation of Cyanine Dyes. Journal of the American Chemical Society. 2015; 137:1073–1080. [PubMed: 25534273]

Yu D, Baird MA, Allen JR, Howe ES, Klassen MP, Reade A, Makhijani K, Song Y, Liu S, Murthy Z, et al. A naturally monomeric infrared fluorescent protein for protein labeling in vivo. Nature Methods. 2015; 12:763–765. [PubMed: 26098020]

Highlights

- Chromophore substitution shifts Arch fluorescence to unprecedented NIR wavelengths
- Directed evolution enhances NIR brightness and affinity for synthetic retinal
- The NIR fluorescence of engineered Arch variants is pH-sensitive
- Membrane-localized Arch variants facilitate NIR imaging in bacterial cell culture

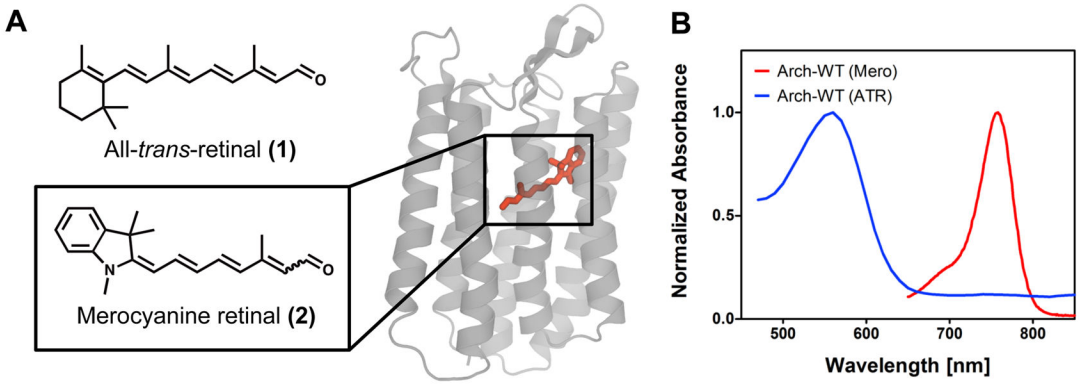


Figure 1. Structure of all-*trans*-retinal and merocyanine retinal

(A) Relative to the natural Arch chromophore, ATR (1), the electron delocalization of merocyanine retinal (2) is extended through the indolylidene ring. Merocyanine retinal was purified in a 3:1 trans:cis ratio at C16. Merocyanine retinal is shown (red) in the retinal binding pocket of an Arch homology model. (B) Merocyanine retinal (Mero) red-shifts the absorbance of wild-type Arch (red curve) by 200 nm compared to ATR-bound Arch (blue curve). Arch-WT = wild-type Arch.

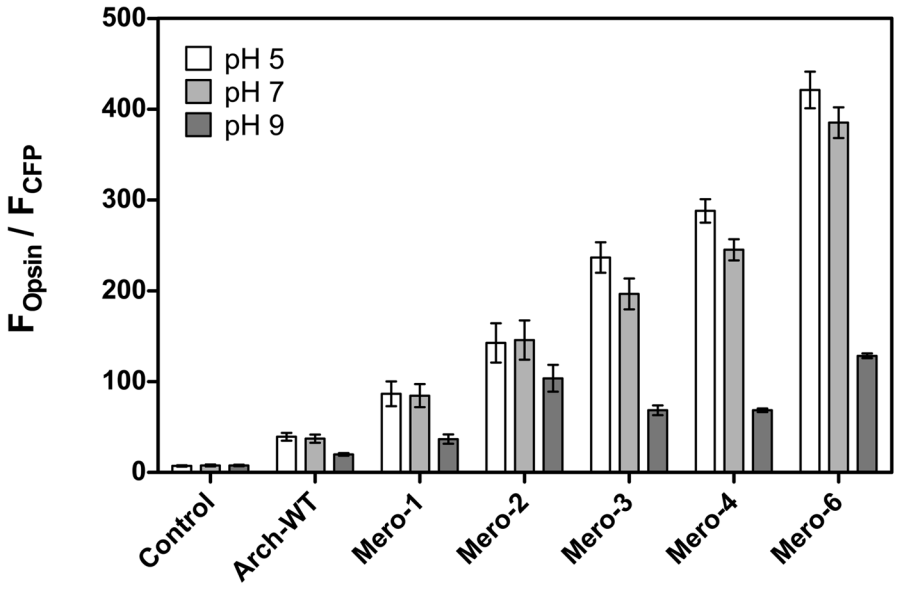


Figure 2. Directed evolution of merocyanine retinal-bound Arch

(A) Expression-normalized NIR fluorescence of merocyanine-bound Arch variants engineered using directed evolution. Fluorescence readings were taken in *E. coli* suspensions with an opsin specific measurement (ex/em 760/785 nm) normalized by the fluorescence of a CFP tag (ex/em 425/475 nm) to account for differences in expression. Cells were grown under screening conditions (1 μM merocyanine added during protein induction) and assayed at pH 5, 7, and 9 (white, light grey, dark grey bars, respectively). Error bars represent one standard deviation of six replicate samples. Mutations in engineered variants: Mero-1 (P60S); Mero-2 (P60S-G61L); Mero-3 (P60S-G61L-P196G); Mero-4 (P60S-G61L-S151A-P196G), Mero-6 (P60S-G61L-T80A-G132S-S151A-P196G).

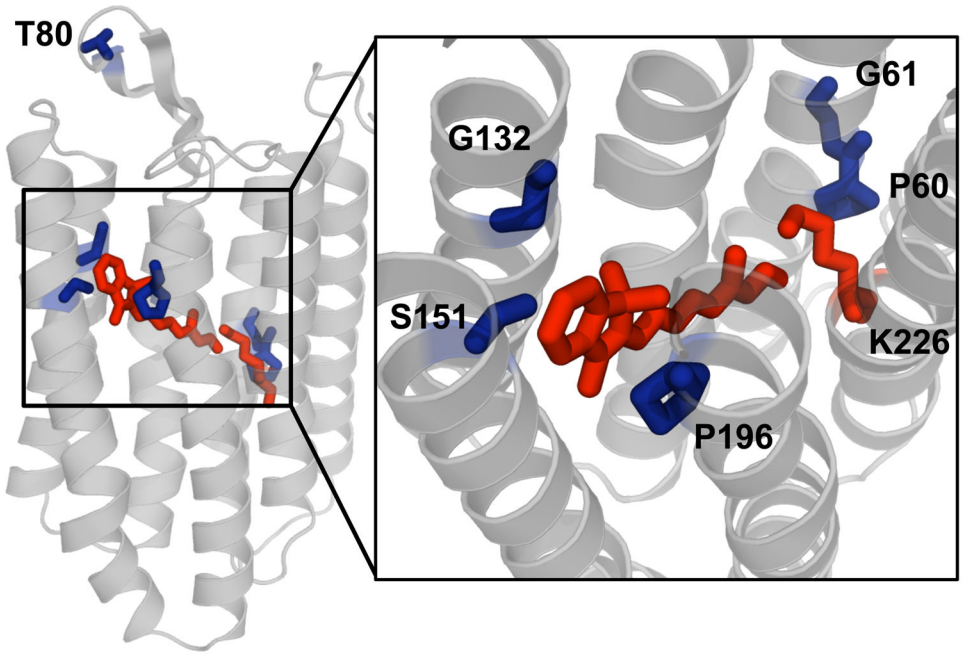


Figure 3. Homology model showing merocyanine retinal in the Arch retinal binding pocket

The Arch homology model is based on the crystal structure of Archaerhodospin-2 (PDB ID: 1VGO). Merocyanine retinal (red) was manually placed over the native chromophore, ATR, to approximate its position within the Arch binding pocket. Although there is space for a Schiff base linkage to the conserved lysine (K226, red), the linkage would require a rotamer change of the lysine which was not assumed for this schematic. Sites yielding beneficial mutations through directed evolution are shown in blue.

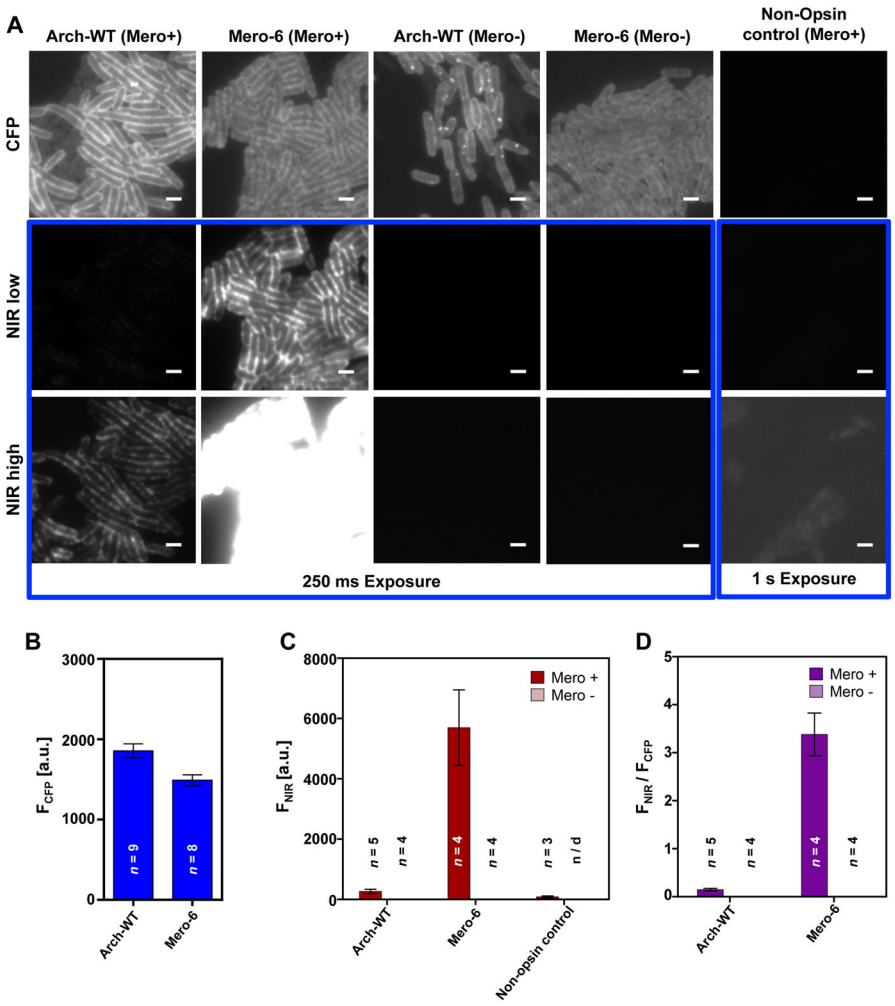


Figure 4. Imaging opsin-specific, merocyanine-dependent NIR fluorescence in live *E. coli.*

(A) Representative NIR (ex/em 727/766–854 nm) and CFP (ex/em 405/>418 nm) fluorescence images of *E. coli* cells expressing wild-type Arch or Mero-6. Low and high contrast images of the NIR channel are shown due to significant differences in the fluorescence of Arch variants; such treatment was not required with the CFP images because expression levels of Arch variants in *E. coli* were similar. *E. coli* samples are expressed with (Mero+) or without (Mero-) 1 μM merocyanine retinal. Scale bar: 2 μm. (B) Mean CFP fluorescence intensity of *E. coli* expressing Arch-WT (n = 9) and Mero-6 (n = 8); each image contains ~50 cells. (C) Mean NIR fluorescence of WT, Mero-6, and control cells expressing a non-opsin protein with merocyanine (n = 5, 4, and 3, respectively) and without merocyanine (n = 4, 4, and n/d, respectively). (D) NIR fluorescence normalized by CFP fluorescence for *E. coli* samples with or without merocyanine added. All error bars represent one standard error of the mean. ‘n’ equals the number of images analyzed.

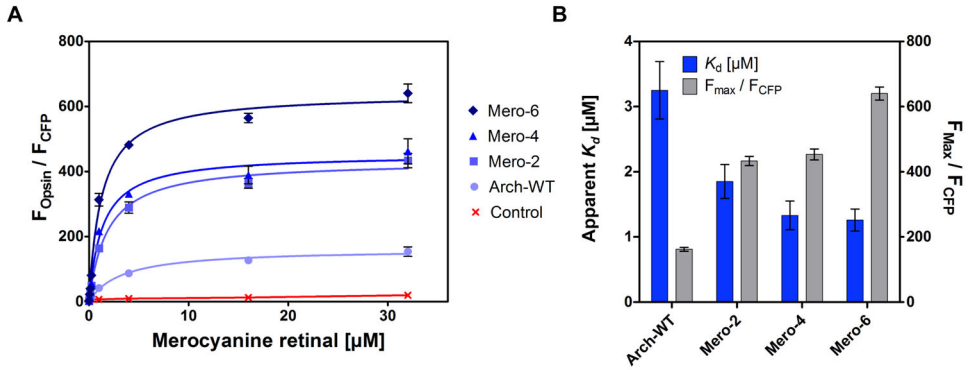


Figure 5. Chromophore concentration-dependent Arch fluorescence in *E. coli*

(A) Normalized fluorescence as a function of merocyanine retinal concentration for control (x), Arch-WT (●), Mero-2 (■), Mero-4 (▲), and Mero-6 (◆). (B) Grey bars: calculated maximal normalized fluorescence of individual Arch variants. Blue bars: calculated apparent K_d of Arch variants (Arch-WT: 3.25 Mero-2: 1.85, Mero-4: 1.33 and Mero-6: 1.26 μM). Error bars represent one standard deviation. Merocyanine retinal had no influence on Arch expression (i.e. the CFP signal; Figure S3*B*).

Table 1

Biochemical characterization of select merocyanine retinal-bound Arch variants and other red-shifted fluorescent proteins.[a]

	Mutations	ATR	Merocyanine retinal						
		K_d (μM)	K_d (μM)	Ex λ_{max}	Em λ_{max}	κ^d	QY	ε (M^{-1} cm^{-1})	Brightness (QY × ε) / 1000
Wild-type Arch	---	8.5	99	757	772	1.79	0.096	78000	7.5
Mero-2	P60S, G61L	11	11	758	775	n/d	0.113	81000	9.2
Mero-4	P60S, G61L, S151A, P196G	26	n/d	761	775	n/d	0.105	82000	8.6
Mero-6	P60S, G61L, T80A, G132S, S151A, P196G	330	17	759	770	1.49	0.130	82000	10.7
Merocyanine retinal [b]	---	n/a	n/a	517	691	n/d	0.12	22000	2.6
mCherry (Kremers et al., 2011)	---	n/a	n/a	587	610	n/a	0.22	72000	15.8
mIFP (Yu et al., 2015)	----	n/a	n/a	683	704	n/a	0.08	82000	6.8
iRFP-720 (Shcherbakova and Verkhusha, 2013)	---	n/a	n/a	702	720	n/a	0.06	96000	5.8
BphP1-FP_C20S (Shcherbakova et al., 2015a)[c]	---	n/a	n/a	639	669	n/a	0.154	105000	16

Ex = excitation, Em = emission, n/a = not available, n/d = not determined, κ = relative exponential fluorescence decay in *E. coli*, QY = Quantum Yield, ε = Extinction Coefficient mEGFP = monomeric Enhanced GFP, mIFP = monomeric Infra-red Fluorescent protein, iRFP = infra-Red Fluorescent Protein, BphP1-FP = Bacterial phytochrome

[a] Buffer for all *in vitro* Arch measurements: 10 mM Tris, 200 mM NaCl, 0.15% DDM (pH 6.5)

[b] aldehyde form of free merocyanine retinal (Figure 1*A*)

[c] bound to the plant bilin chromophore, phytochromobilin

[d] To correct for a non-uniform laser intensity profile, the exponential decay rate of the NIR protein is shown relative to the exponential decay rate of the fused CFP. Fluorescence measurements in the respective channels were normalized for measured laser intensity and the absorption profile of the fluorophore. Thus a κ value of 1 indicates that the NIR fluorescence decays at the same rate as CFP (0.1 s^{-1}, where 36.8% of the fluorescence decays in 10s) under comparable light intensity and wavelength. (See Figure S4 for analysis and the fluorescence decay rates of each channel)